II. STATE-OF-THE-PRACTICE OF ITS DATA ARCHIVING AND USES

The information presented in this section will provide the reader with insight into the state-of-the-practice in both the archiving practices at existing TMCs and the uses of ITS data. First, select experiences of data sources and related data archiving and management practices throughout North America are investigated. A majority of these experiences are drawn from traffic management centers (TMCs) where data are reported to the centers at fixed intervals. Discussion is then provided that explains the data source provided by commercial vehicle operations (CVO) at weigh stations. Data for CVO are often collected at weigh stations and are not reported to a TMC; however, CVO data collection is discussed in this section due to the potentially rich data source that can be obtained. Finally, current activities using ITS data are discussed. Appendix A shows the contacts from which the information in section II of this report was obtained.

A. Archiving Practices at Existing Transportation Management Centers and Operation Centers

Recent research has been performed that evaluated the current practices of retaining ITS data for future transportation analyses (1,2). The research found that the uses and users of ITS data appeared to be relatively consistent at the various TMCs. They include requests for traffic counts to estimate construction impacts, feeds to the Internet or media for real-time reporting of the data, requests for planning and modeling uses, and researcher requests. While the different TMCs have similar uses and user requests, the aggregation and storage practices are not consistent. Of the 15 TMCs interviewed at the time, 20 percent were not saving any of the ITS data. Aggregating and saving the data at 1 minute or less was practiced at 60 percent of the TMCs. The remaining 20 percent of TMCs save the incoming data at an aggregation level between 5 and 15 minutes.

As part of this research effort, the research team re-evaluated the list of TMCs from this previous research to identify any changes in data archiving or management strategies. Generally, it was found that the TMCs operators and managers have recognized the value of the ITS data for secondary uses and are interested in saving, formatting, and aggregating data to facilitate other uses beyond the traditional real-time needs of the operation center. There appears to be a stronger interest in interdisciplinary coordination (i.e., operations personnel working more closely with planning personnel) although financial constraints are often the leading hindrance to such efforts. The research team was often told that when budgetary constraints are being considered, it is often the data collection and/or management areas that are reduced. This section will describe data archiving and retention practices at select TMCs. Finally, this section of the report concludes with additional observations and lessons learned.

Toronto's COMPASS

COMPASS is a traffic management system that monitors portions of Highway 401 in Toronto, Canada. The center monitors over 2,800 loops as well as numerous CCTV cameras. The loop detector data are polled every 20 seconds and report volume, occupancy, and an average speed from an assumed vehicle length for the single-loop detector stations. Whenever maintenance is required of the highway asphalt, double detectors are being installed to replace the single loop detectors.

The 20-second data are aggregated to the five-minute, 15-minute, one hour, daily, and monthly time periods. The TMC archives all data for 20-second and 5-minute time increments. For data summaries of 15 minutes or more, only volume data are saved. The five-minute time increment was selected because it appeared to provide a convenient time increment for many users. Common users and uses of the data include input to the real-time incident detection COMPASS algorithm, Internet flow maps, in-house requests, and researchers desiring 20-second data for simulation and algorithms. In-house request account for about 60 percent of data request (e.g., traffic forecasts, roadway impact analyses).

The data have traditionally been archived on 8 mm tape cartridges. The tapes require special equipment and software needs to download the data, and it was found that this system was relatively difficult for obtaining historical data. Further, over extended periods of time (i.e., greater than about 5 years), the tapes begin to lose data and reliability.

In response to these concerns with data archived on tape, since about 1997, the COMPASS system has been saving the data to CD, and they have been developing compatible software for obtaining the data via CD. The CD itself is being developed to contain its own software that will create and display summaries of calculations in table form of interest to the user. COMPASS personnel are currently assessing the appropriate and most useful format and aggregation level for the CDs for a variety of users and uses (i.e., should the CD contain only 20-second or hourly data). This decision will likely be made depending upon the demand for the CDs that are being developed at the different aggregation levels. One other difficulty of the CD application is that it operates slowly.

This is a relatively new effort with regard to data archiving at the COMPASS TMC, and the TMC personnel are in the process of developing their policies regarding sharing the data. The experience at the COMPASS system illustrates the interest in discovering a means to allow more efficient storage and retrieval of available ITS data.

Michigan ITS Center

The Michigan ITS (MITS) Center provides an example of another TMC that is developing new standards on an aggregation level and determining an on-line data access strategy. The Center contains two loop data collection systems. The older system covers 52 km (32 miles) of freeway with single loops located at 0.54 km (0.33 mile) spacing. The data are sent to the Center every 10 milliseconds and aggregated to hourly lane volumes which are saved on tape backup. The newer

system covers 240 km (150 miles) of freeway with dual loops at a 3.2 km (2 mile) spacing. These data are aggregated in the field at 20-second intervals and then sent to the Center. These data are then aggregated up to 1-minute and the volume, occupancy, and speed are saved. Personnel at the Center hope to eventually fill in additional dual-loop detectors between the existing loops in the newer system. The Center keeps one week of data on-line at the 1-minute aggregation level. Data older than one week are saved to tape backup.

As with many areas, personnel at the MITS Center are interested in developing better access to archived data for a variety of users and uses. They are evaluating the possibility of a common aggregation level to satisfy all users as well as trying to develop a more efficient on-line data access. Personnel at the Center commented that this is difficult since different users seem to inevitably require different types of data at different aggregation levels. They would also like to incorporate their changeable message sign (CMS) log into this data set. The CMS log includes information such as when a message went up, was removed, and what was written for a given incident condition. Personnel at the Center also expressed the desire to integrate information that comes into the Center through public telephone and the highway advisory radio (HAR) into the database. They will likely consider CD or digital video disk (DVD) for archiving rather than the tapes that are currently used.

Another concern at MITS are loop failures. In the hourly volume summaries that are saved, a measure of the lane operability is also saved. The measure is the loop-minutes of failure which is the number of loops failing times the number of minutes failing. When making temporal or spatial comparisons, this measure provides some insight into the extent to which a loop(s) may not be reporting. For example, if the loop operability equals nine, this could mean that three loops are out for three minutes, or one loop is out for nine minutes. This does not evaluate the extent that volumes may be reading too high or too low, and these data quality issues are also a concern at the Center.

Oak Park, Illinois Traffic Systems Center

The Illinois Traffic Systems Center is another TMC within which archiving and data retrieval issues are being strongly re-evaluated. The Center is currently obtaining volume and occupancy and estimating speeds from the single-loop detectors in the field. The 20-second data are aggregated to the 5-minute level, and only the occupancy data are saved at this level. The occupancy measures are saved since this is the measure that is often used for freeway management at the Center. Hourly volume data are also saved. Both the 5-minute and hourly data are saved to tape.

The Center personnel have an interest in using the available data for many secondary uses and want to be able to support the demand. They have recently hired a consultant to evaluate the data retention and archiving in the Center. Many issues are being considered in this effort including determining what should be saved, at what level should it be saved, what medium should be used for storage, and what type of automation of such a system can be developed. It is anticipated that about one year of data is desired on-line, and a database management system that saves the data to CD, rather than tapes, is also desired. Sybase will be used for the database, and one of the first tasks of the consultant is to evaluate the potential uses and users of the data.

Maricopa County, Arizona

The Traffic Operations Center (TOC) in Phoenix is used for many purposes including freeway management, and it began operations in September 1995. All the 20-second loop data that comes into the Center are saved. The loop data includes volume, occupancy, and an estimate of speed from the single loop detectors in the field. Five minute summaries of the data are also created and saved. Past telephone interviews with TOC personnel indicated that data storage may become a significant concern, and due to the large size of data files used for analysis, it is often difficult to manage the data sets with most spreadsheet and database software.

Efforts at the Maricopa County Department of Transportation (MCDOT), with the Arizona Department of Transportation (ADOT), are developing a system to meet data archiving and access needs (3). The system is called the AZTech Data Server (ADS), and is shown in Figure 1. The formatted records one through six at the top of the figure represent data inputs from different sources such as incident data or loop data that come into the Advanced Traffic Management System (ATMS). The data spooler/logger represents the element of the system that provides the quality control and screening of the data as it is received from the different sources and then logs the data. Data are then provided to ETAK, which is currently developing their traveler information service to the public, the ADS, and to other agencies that have a feed. As shown in the figure, it is within the ADS that user-defined queries of the data can be performed for both on-line and off-line applications.

While loop and incident data are currently the primary data inputs, it is anticipated that many other sources will also become available. The objective of the project is to, "enhance the ADS to provide and maintain valid, classified ITS-derived data for use in planning, modeling, and real-time operation applications." Ultimately, the ADS would provide an interface for all data types that enter the system. The project is funded at $500,000 through congestion mitigation and air quality (CMAQ) funds and will end September 2000. The project is being managed by MCDOT. A Project Manager within the Transportation Planning Division will manage the project, a Technical Working Group (TWG) will be assembled from interested jurisdictions, and additional input will be provided to the project from regional ITS coordination committees.

The first steps of the project are underway. These include evaluating available user services, data screening and formatting issues, and appropriate validation and quality control checks. It is anticipated that different users will accept differing levels of data validity, and the system developers are considering such rule-based methods for queries.

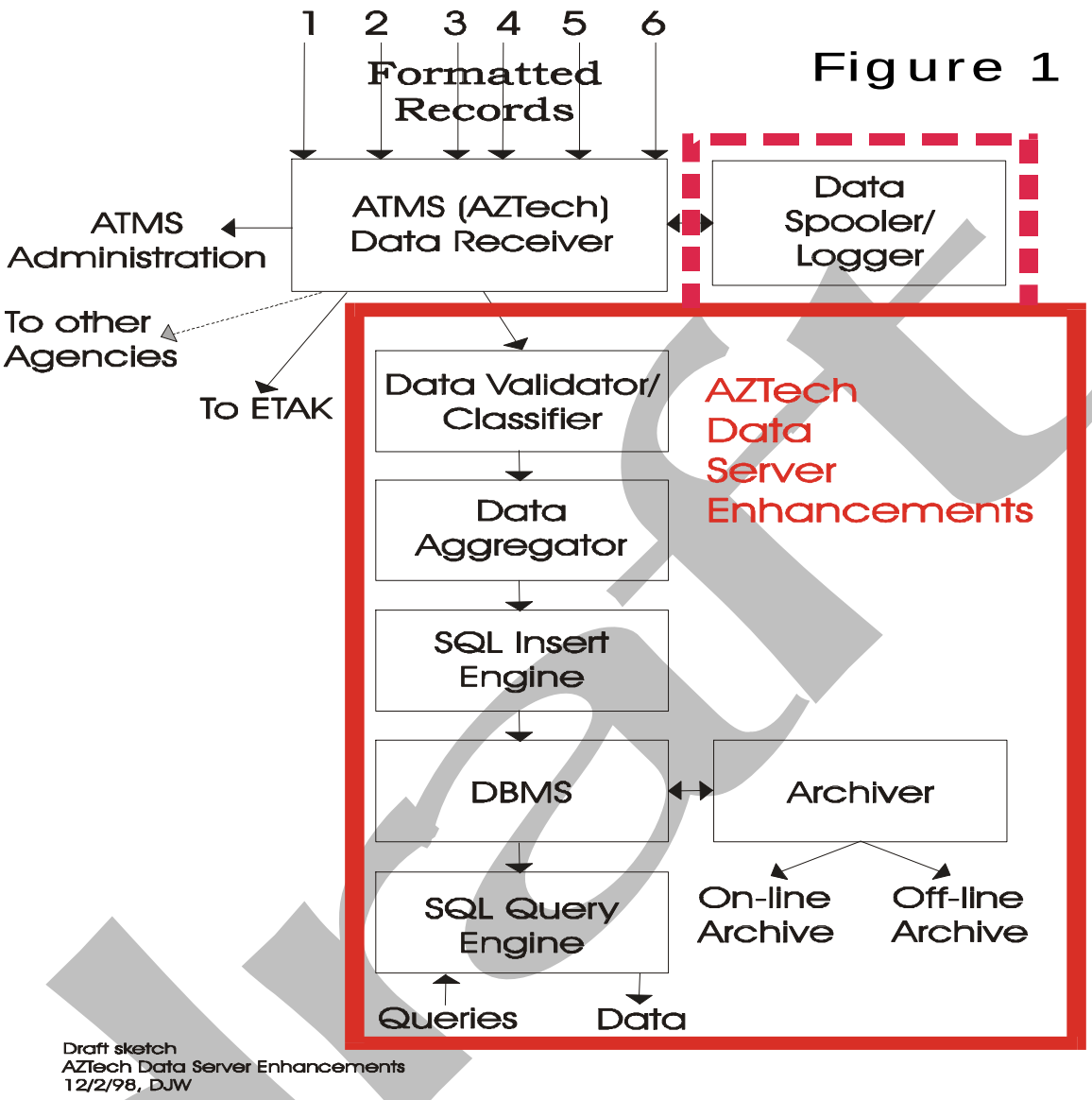

Figure 1. Illustration of the AZTech Data Server (Adapted from Reference 3)

Another effort underway at the Maricopa Association of Governments is being performed as part of a congestion study in the region. Speed, volume, and classification data were desired on freeways and arterial roadways. Aerial photographs were used to obtain density on freeways and queue lengths on arterials. Volume data were obtained from AZTech. Video was used to obtain classification data. Fifteen minute volumes were recorded from the video data, and it was compared to the AZTech volume data. One year of data at 5-minute aggregation was supplied by AZTech. It was found that sometimes the data matched well, and other times the AZTech loop detectors were not operating or the detector volumes were low. Personnel at the Association indicated that it was sometimes difficult to identify where the loop detectors were in the field for the data that was being received at the Center. The Association has been in contact with AZTech about these data quality issues and was informed that AZTech would take efforts to maintain the loop detectors. Personnel at the Association are leery of the quality of data being collected by the loop detectors at the Center.

Los Angeles District 7 Traffic Management Center

The Los Angeles District 7 TMC is in the process of developing a new system for data archiving and access. The TMC is primarily used to monitor and verify incidents in the freeway coverage area. Both inductance loops and closed circuit television (CCTV) are in operation in the monitoring system. The TMC polls the loop stations every 30 seconds. Prior to the development of the new archiving system, the TMC would save three days of 30-second data and four days of five minute summaries into temporary storage. The 30-second data were all saved to tape since the TMC was opened. The DOT had developed special software that was necessary for data retrieval from the mainframe computer.

The new system will provide instant access to 13 months of on-line data through an Oracle database for personnel in the Center. Beyond the 13-month time period, the data are currently backed up on tape. In the future, as part of the new system development, it is anticipated that backing up to tape will be replaced by archiving the data to CD. Future plans also include evaluating archiving different aggregation levels to a given CD (e.g., having a CD for daily summaries, one for 15-minute data, etc.). This CD system would act as a CD clearinghouse for data for interested parties. The system being developed in Los Angeles is the prototype development of the system that, if successful, would be implemented in all Caltrans TMCs.

Minnesota Traffic Management Center

The Minnesota Traffic Management Center (TMC) monitors over 75 percent of the freeways in the twin cities metropolitan area. The center coverage expands about 10 percent per year. The most significant source of ITS data in the Minnesota TMC are the data from the 3,800 loop detectors. The detectors provide volume, occupancy, and an estimate of speed based upon an assumed vehicle length. These data are obtained every 30 seconds, and every five minutes the data are aggregated and saved. Daily information is stored to a CD and about 6 months of data can be saved to one CD. The 30-second data are also saved. The process is automatic and the data have been saved since the center opened in 1993.

Most requests for the data are from the DOT for traffic analysis, construction impact determination, and planning applications. Researchers from local universities often request the data as well. Another significant use of the data is for the evaluation of ramp metering strategies. This use has fostered an interest in more detailed study and calibration of the loops. This effort is explained in greater detail in the next section of this report describing current uses of ITS data.

Montgomery County Transportation Management Center

The TMC in Montgomery County, Maryland has also been the location of activities related to data archiving for secondary uses of ITS data. Personnel at the Center save the1-minute volume and speed data that are obtained from loop detectors to the 5-minute aggregation level. The loop detectors are located on arterial streets on the downstream legs from the intersections. This information is used to assist in determining offsets and splits for signal timings along the arterials.

The Center personnel also have an increased interest in developing a more strategic archiving plan for the data. They are currently working on a data archiving effort funded at $100,000 and shared by several public and private sector participants to assist in determining elements such as identifying users and uses of the data, a common-denominator of data aggregation to satisfy these uses, and a method or system to query data of interest about a particular link. An application of this effort is also discussed in greater detail in the next section of this report.

Summary of Transportation Management Center ITS Data Archiving Activities

Several observations and lessons learned were noted during the conversations with TMC personnel. These observations include the following:

Improved Interdisciplinary Coordination: It was noted that many of the TMC personnel that are involved with real-time traffic management activities are more interested in supporting the data needs of secondary uses. For example, there appears to be increased support by operations personnel to make data accessible and in the correct format for planning applications. Although many agencies are still developing these formats, storage media, and interfaces, the coordination has improved.

Seeking Alternative Funding Sources: In previous discussions, it was clear that there was interest by various groups in developing improved archiving practices, yet funding was limited or not available. Recent interviews revealed that agencies are pursuing more "innovative" funding opportunities rather than using money from the traditionally itemized data collection and related activities. Funding sources mentioned included the use of congestion mitigation and air quality (CMAQ) funds or State Planning and Research Funds.

Similar Concerns Among Practitioners: There was a lot of interest and support in the work being performed and the questions being raised during the telephone conversations. Individuals were happy to provide the information and were eager to find out what other areas were doing to address the concerns. Nearly all of the experiences above indicate similar interests and concerns in different regions. These questions included the common issues of, who are the users of the data, what are the data uses, what data should be kept, at what aggregation level should it be aggregated, what type of

interface can be developed, how can the system be automated, what is the appropriate data format, and what data elements should be included in the database. In addition, concerns were expressed related to the learning curve of the computer systems (e.g., Oracle, Sybase) by both operations personnel and some secondary users (e.g., planning organization personnel) since these individuals are not used to operating such systems.

Data Quality Concerns: Many individuals noted concern over the quality of the data itself. Although there is an increased interest in archiving practices, there is also an increased data quality concern. Many expressed that more work needs to be performed in identifying the quality of the data itself. In addition, it was often noted that referencing the actual data obtained from the TMC with a location on a roadway segment was occasionally difficult.

Improved Data Access: It is interesting to note that although many of the TMCs do not have fully-developed data warehouses or archiving practices, they are making significant progress in making the data more accessible. Incrementally, improving data access is a great improvement. In fact, there are few TMCs that have data that are relatively easy to access. One example of how data access is being facilitated in many areas is through the improvement of storing data to CD rather than tape. Finally, it appears that once data access itself is improved, many secondary data users discover or develop appropriate software to produce the results that meet their needs.

REFERENCES

1. R.E. Brydia, S.M. Turner, W.L. Eisele, and J.C. Liu. "Development of Intelligent Transportation System Data Managment." In *Transportation Research Record 1625*. Transportation Research Board. Washington, D.C. 1998.

2. S.M. Turner, R.E. Brydia, J.C. Liu, and W.L. Eisele. *ITS Data Management System: Year One Activities.* Texas Transportation Institute, Research Report 1752-2. College Station, Texas. August 1997.

3. Wolfson, D. *"Scalable AZTech Data Server Enhancements for Planning and Operations."* Draft Scope of Work. Maricopa County DOT. March 1999.

APPENDIX A

Contacts for information provided in section II. A. and II. B.

David Tsui
COMPASS
Ontario Ministry of Transportation
Toronto, Ontario, Canada
Phone: (416) 235-3538

James Aswegan
Minnesota TMC
Minnesota DOT
Minneapolis, Minnesota
Phone: (612) 373-2761

Tony Cioffi
Traffic Systems Center
Illinois DOT
Oak Park, Illinois
(708) 524-2145

David Lau
Los Angeles District 7 TMC
Caltrans
Los Angeles, California
Phone: (213) 897-0341

David Wolfson
Maricopa County DOT TMC
Transportation Planning Division
Maricopa County, Arizona
Phone: (602) 506-6950

Mark Schlappi
Maricopa Association of Governments
Maricopa County, Arizona
Phone: (602) 254-6300

Bruce Mangum
Montgomery County TMC
Montgomery County Department of Public Works and Transportation
Montgomery County, Maryland
Phone: (301) 217-2197

Don Ostrander
Maryland National Capital Park and Planning Commission
Montgomery County, Maryland
Phone: (301) 495-2184

Arvyd Satraitis
Michigan ITS Center
Michigan DOT
Detroit, Michigan
(313) 256-9800

Jennifer Walton
Kentucky Transportation Center
Lexington, Kentucky
Phone: (606) 257-4724